

Testing Müllerian mimicry: an experiment with wild birds

Michael P. Speed^{1*}, Nichola J. Alderson¹, Christine Hardman¹
and Graeme D. Ruxton²

¹*Environmental and Biological Studies, Liverpool Hope University College, Hope Park, Liverpool L16 9JD, UK*

²*Institute of Biomedical and Life Sciences, Division of Environmental and Evolutionary Biology, Graham Kerr Building, University of Glasgow, Glasgow G12 8QQ, UK*

Experiments with wild birds feeding on pastry 'prey' were performed to test competing theories of Müllerian mimicry. Conventional theories predict that all resemblances between defended prey will be mutually advantageous and, hence, Müllerian. In contrast, unconventional theories predict that, if there are inequalities in defences between mimetic species, the less well-defended prey may dilute the protection of the better defended species in a quasi-Batesian manner. This unconventional prediction follows from an assumption that birds learn about the edibilities of prey using rules of Pavlovian learning. We report on two experiments, each lasting 40 days, which showed that a moderately defended prey can dilute the protection of a better defended mimic in a quasi-Batesian fashion, but can add protection to a mimic which has the same moderate levels of defence. These results match predictions of unconventional theories of mimicry and go some way to resolving the long-running arguments over the nature of Müllerian mimicry.

Keywords: warning signal; mimicry; predator; evolution; Pavlovian learning

1. INTRODUCTION

Despite many years of study (Müller 1879; Turner 1987; Huheey 1988; Mallet & Joron 1999) and the development of extensive theory (Müller 1879; Huheey 1976; Owen & Owen 1984; Turner *et al.* 1984; Endler 1991; Speed 1993a,b, 1999; Turner & Speed 1996; MacDougall & Dawkins 1998; Speed & Turner 1999a,b), Müllerian mimicry remains a largely untested idea. The consequent absence of data has caused a long-running controversy (Turner 1987; Ruxton 1998; Speed 1998; Mallet & Joron 1999) about the nature of Müllerian relationships and has left a key example of Darwinian adaptation poorly understood. At the heart of the controversy is a behavioural question about the way that predators deal with differences in the levels of chemical (or other) defences between mimetic prey (Huheey 1976; Owen & Owen 1984; Turner *et al.* 1984; Speed 1993a).

In his original formulation, Müller (1879) (see also Joron & Mallet 1998; Mallet & Joron 1999) assumed that predators need to attack a fixed number of defended, inedible prey before learning to avoid them completely. He then proposed that mimicry between two defended species will be beneficial to both because the fixed costs of predator education will be shared between a larger population of prey. Müllerian mimics thus have strength in numbers and new, rare polymorphic forms will not be favoured in such a system (Turner 1987). This enables a further important prediction that all mimicry between defended species should be monomorphic. The conventional interpretation is thus one of 'universal monomorphic Müllerian mimicry' between defended species, even if there are differences in defensive properties between prey (Turner *et al.* 1984; Turner 1987; Endler 1991). This view has received renewed support (Joron &

Mallet 1998; Mallet & Joron 1999), despite a lack of direct empirical evidence and the existence of polymorphism in a number of defended mimetic species.

In contrast, unconventional theories of mimicry (Huheey 1976; Owen & Owen 1984; Speed 1993a, 1999; MacDougall & Dawkins 1998; Speed & Turner 1999a,b) predict that predators may continue attacks on moderately defended prey in times of nutritional need, because they are prepared to trade off the benefits of nutrition against the costs of toxins (Speed 1993b). This matches a view of behaviour based on models of Pavlovian learning (Speed 1993a). In these theories, mimicry by moderately defended prey species of better defended species may not therefore be mutualistic and Müllerian but instead may be parasitic and quasi-Batesian. Quasi-Batesian mimics are predicted to have highest fitness when rare (though see Speed 1999). This strength in rarity should favour new, rare mimetic polymorphisms and may explain the otherwise troubling presence of polymorphism in some aposematic mimics (Huheey 1976; Owen & Owen 1984; Speed 1993a; Speed & Turner 1999b). In different circumstances the same moderately defended species may be Müllerian if, for example, it resembled another moderately defended species or a better defended but rare co-mimic (Speed 1993a, 1999). However, such moderately defended species could also function as models to truly edible Batesian mimics, particularly in multispecies systems (Speed & Turner 1999b). The role of a defended mimic in unconventional theory is thus highly sensitive to ecological context. This is quite different to the traditional dichotomy of classical Batesian and Müllerian mimicries (e.g. Turner 1987).

Since differences in edibility between defended species are well known (Brower *et al.* 1963; Sargent 1995) and since mimicry is an important case study in Darwinian adaptation (Turner 1987; Mallet & Joron 1999), experimental tests are clearly needed to resolve the controversy.

*Author for correspondence (speedm@hope.ac.uk).

We report experiments which, for the first time to the authors' knowledge, test key predictions of conventional and unconventional theories of Müllerian mimicry. Using pastry baits as prey and wild garden birds (predominantly blackbirds, with some sparrows, robins and starlings) as predators, we examined the effects on predation when a moderately defended pastry 'prey' mimicked a better defended 'prey' and also varied its density. Conventional Müllerian theory predicts a mutualistic relationship, whereas unconventional mimicry theories predict a parasitic relationship, in which protection of both co-mimics declines if the density of the less well-defended species increases. We also investigated whether mutualistic Müllerian mimicry can be demonstrated when there is mimicry between moderately defended prey. The experiments reported represent the first test of opposing mimicry theories and demonstrate the importance of data in mimicry studies.

2. METHODS

(a) *Artificial prey*

The general methods were from Morrell & Turner (1970) (and also Marples 1993). We made artificial prey from pastry (335 g flour, 135 g lard and 30 g water) which was dyed with 10 ml of food colouring to each 500 g of pastry. The prey were each 2.5 cm in length and 1 cm in diameter and mounted on equilateral triangles of coloured card (3 cm in height) secured with a blunted drawing pin pointing upward from the underside of the card. The prey were assigned to one of three levels of acceptability: high defence (1.25 g quinine hydrochloride and 2.5 g of mustard per 500 g of pastry), moderate defence (0.5 g quinine hydrochloride and 1 g of mustard per 500 g of pastry) or edible (no quinine or mustard).

There were five types of prey used in the experiment (details in figure 1). The model and mimic had the same bait and card colours as each other, but the model was highly defended and the mimic moderately defended (the terms model and mimic are used for convenience and are not intended to pre-judge mimetic relationships). The model and mimic had controls (model control and mimic control) which were non-mimetic prey with unique appearances (figure 1), but which had levels of defence which matched those of the model and mimic, respectively (figure 1). They therefore served as controls for mimicry, showing how the model and mimic would be attacked if they shared no resemblance. In addition, an edible undefended prey called the edible control, again of unique appearance, was used to monitor the demand for food (figure 1).

The cards were attached to the pastry in order to (i) make all prey types conspicuous, (ii) generate visual gestalts (Ikin & Turner 1972) which would aid identification, and (iii) partly control for the colour preferences of wild birds (Lea & Turner 1971; Marples *et al.* 1998). Hence, if the birds in question were, for example, disinclined or overinclined to sample yellow prey, then we might see common predation levels in all items containing yellow (palatable control, model and mimic; figure 1) and similarly for red and blue (model control and mimic control; figure 1). Generalization between 'non-mimetic' prey which shared colours does not seem to have affected our results.

(b) *Experimental design*

Two experiments were carried out in different locations in Liverpool, each on an area of lawn measuring 3 m × 3 m.






prey type	bait/card colours	edibility	daily numbers:	
			condition A:	condition A:
			● mimic common	● mimic rare
			● mimic control rare	● mimic control common
edible control	yellow bait green card	edible	20	20
				
model	green bait yellow card	highly defended	20	20
				
mimic	green bait yellow card	moderately defended	20	5
				
model control	blue bait red card	highly defended	20	20
				
mimic control	red bait blue card	moderately defended	5	20
				

Figure 1. Summary of prey construction and numbers of each put out each day. Only the mimic and mimic control vary densities between conditions A and B which both lasted for 20 days. In condition A, the mimic is common (20 presented each day) and the mimic control is rare (five presented each day). In condition B, the mimic is rare (five presented each day) and the mimic control is common (20 presented each day). Conventional theories predict (i) that the moderately defended mimic should add protection to the model, particularly when the mimic is common, whereas unconventional theories predict that the moderately defended mimic should dilute the model's protection. Both conventional and unconventional theories predict (ii) that the mimic should gain protection from mimicry. However, (iii) while conventional theory predicts that the mimic's gain should be highest when it is common, unconventional theory predicts that it should be highest when rare. Both types of theory predict (iv) that the moderately defended mimic control which is not involved in mimicry should gain highest protection when common, and (v) that the highly defended model control which is also non-mimetic should be attacked at a constant, low rate.

Experiment 1 was conducted between 3 February and 14 March 1995. Experiment 2 was conducted between 9 January 1996 and 17 February 1996. The perimeter of the experimental arena was marked by wooden stakes, which divided each side into 30 cm sections and provided 100 squares in the arena on which we randomly positioned the artificial prey (see Morrell & Turner 1970). Both experiments were preceded by four weeks of feeding with bread to encourage foraging in the area, followed by ten days pre-experimental familiarization during which the prey were presented (20 samples of all prey except the mimic, of which five were presented).

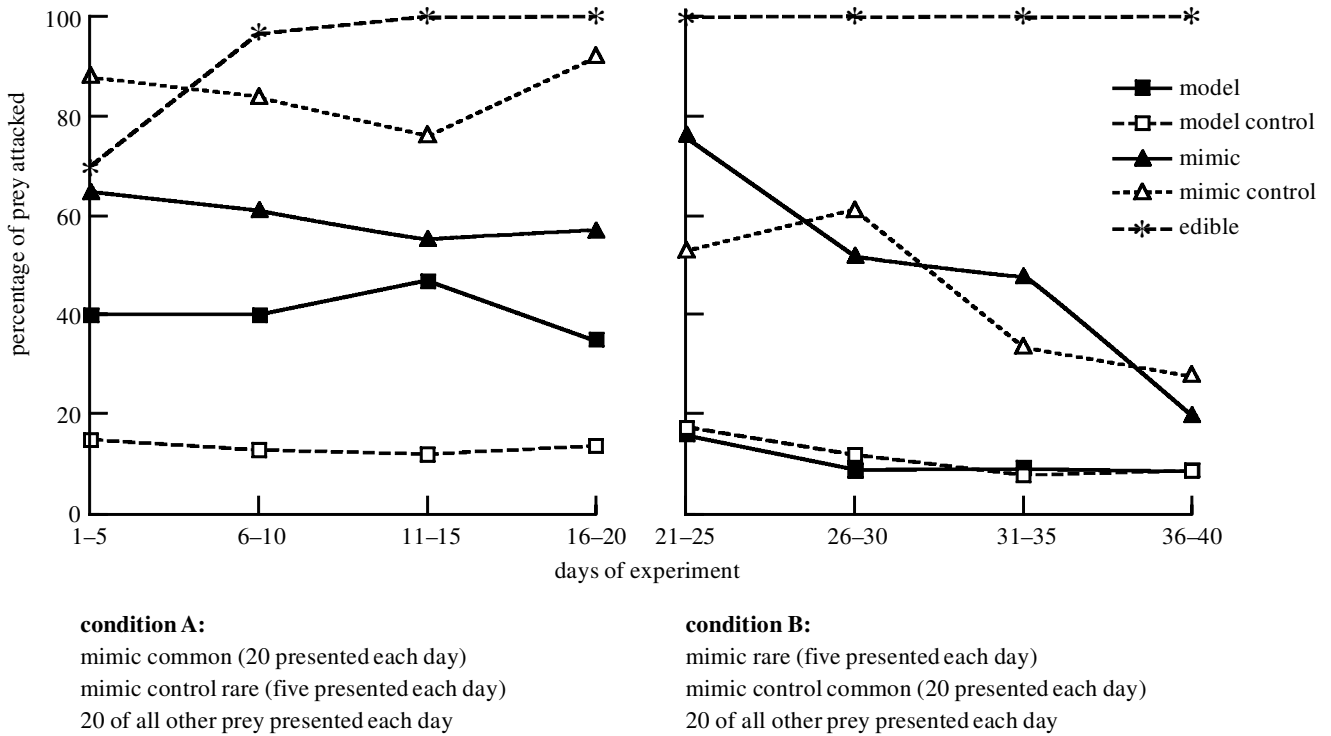


Figure 2. Attacks on mimetic and non-mimetic prey: experiment 1. The percentage of each prey type attacked is shown on the y-axis and time in 5-day periods on the x-axis. In this experiment mimic is common for the first 20 days (condition A) and rare during days 21–40 (condition B). Conversely mimic control is rare for the first 20 days and common during days 21–40. All other prey are common throughout the experiment (i.e. 20 presented each day).

The experiments both ran for 40 days. The daily numbers of the edible control, model and model control presented were constant at 20 of each. The numbers of the mimic and mimic control were varied within an experiment according to one of two prey conditions which each lasted for 20 days (figure 1). In condition A, the mimic was common (20 prey each day) and the mimic control rare (five prey each day), whereas in condition B the mimic was rare (five prey each day) and the mimic control common (20 prey each day). The total numbers of prey put out each day were therefore constant at 85 throughout the experiment (figure 1). This design enabled us to contrast the predation rates on the model and mimic when the mimic was common (condition A) or rare (condition B). It also enabled us to consider how changes in the mimic control's density would affect the frequency with which it was attacked. If the birds judged the mimic control to be defended (as we intended), then the proportion attacked should be lower when the mimic control is common (condition B) than when it is rare (condition A).

In experiment 1, the order of conditions was condition A, the mimic common and then condition B, the mimic rare, whereas in experiment 2 the order was reversed. Prey were put out at 08.30 and collected at 16.00. Birds were observed feeding for approximately 1 h each day throughout both experiments. We scored whether baits were pecked or removed ('eaten'). 'Pecked' prey can, to a limited extent, be classified as pastry which has been rejected before ingestion. Prey removed from the arena may have been eaten by visiting birds or, alternatively, may have been removed from the garden without ingestion. Blackbirds in particular were sometimes seen picking up bait and flying off with them to places of shelter. For the most part then, we pooled counts of pecks and of removals into a single category—the

number of prey attacked. When analyses had been performed on the proportions of prey presented which were attacked the arcsine square-root transform was used.

3. RESULTS

(a) Predator learning and prey edibility

The results are shown in figures 2 and 3 as the percentages of prey attacked in sequential blocks of five days. There is some evidence of predator learning during the experiments. During the first ten days of experiments 1 and 2, the percentages of the edible controls attacked were 84 and 76, respectively, but in the remaining 30 days all the edible control prey were attacked (the vast majority being classed as eaten). In experiment 2, the first ten days are clearly a time for predator recruitment and learning. We have not excluded this data from the analyses but do report the rare instances where inclusion of this learning period makes a difference to the interpretation of the data.

Clearly the birds found the edible control completely edible and the well-defended model control highly aversive (particularly in experiment 1) (figures 2 and 3). The mimic control, which contained lower levels of aversants, was attacked at intermediate levels when common (figure 2, days 21–40 and figure 3, days 1–20). The birds did therefore discriminate the prey into three levels of acceptability.

(b) The effect of mimicry on the model

To compare the attack rate of the model relative to its control, we calculated the number of models attacked on a given day divided by the number of models plus their

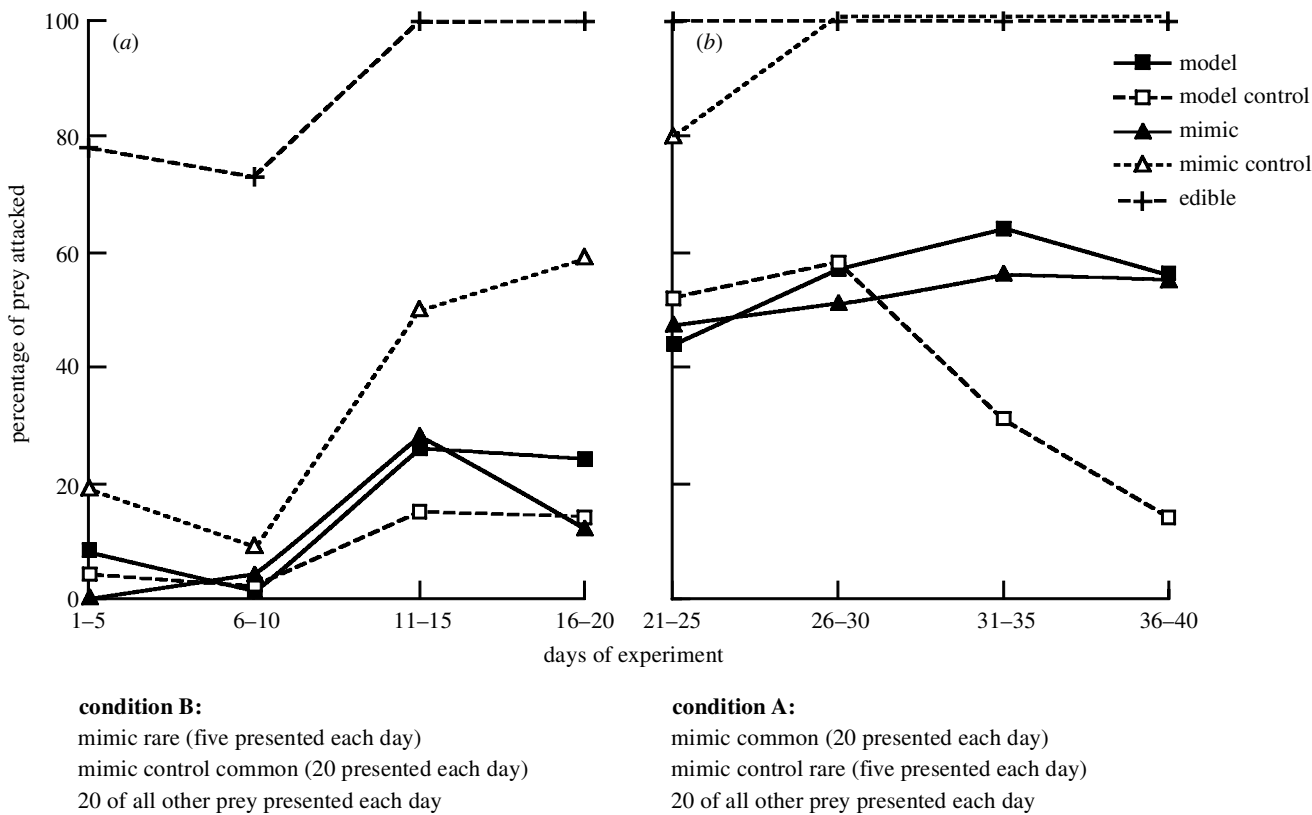


Figure 3. Attacks on mimetic and non-mimetic prey: experiment 2. The percentage of each prey type attacked is shown on the y-axis and time in 5-day periods on the x-axis. In this experiment mimic is rare for the first 20 days (condition B) and common during days 21–40 (condition A). Conversely mimic control is common for the first 20 days and rare during days 21–40. All other prey are common throughout the experiment (i.e. 20 presented each day).

controls attacked that day. This gave a value of between 0 and 1 for all but five days when no items of either type were attacked. The higher this value is, the more the model is attacked relative to its control. For the 75 days for which this quotient could be calculated, we fitted a general linear model (SPSS, v.7) with experiment (1 or 2) and design (condition A or B) as fixed factors and day within a block (days 1–20) as a covariate. This analysis showed no effect of the experiment ($F_{1,70} = 2.25$ and $p > 0.1$), but effects of design ($F_{1,70} = 12.3$ and $p < 0.001$) and day within a block ($F_{1,70} = 8.4$ and $p < 0.01$). Visual examination of the data revealed that the effect of design was that the calculated quotient was higher under design A (condition A) than design B (condition B) and that it increased over time within a given block (this is particularly true in experiment 2).

We also repeated the same analysis simply using the total number of models attacked on a given day as the dependent variable. This time we did find a difference between experiments 1 and 2 ($F_{1,75} = 9.1$ and $p < 0.01$); hence, we analysed the two experiments separately. Experiment 1 showed an effect of design ($F_{1,37} = 101.4$ and $p < 0.001$) but not day within a block ($F_{1,37} = 0.92$ and $p > 0.3$). Experiment 2 showed an effect of design ($F_{1,37} = 61.6$ and $p < 0.001$) and day within a block ($F_{1,37} = 7.0$ and $p < 0.05$). Visual inspection of the data revealed that the effect of design in both experiments was that more models were attacked each day under design A than under design B.

Taken together, these results show that the model suffered higher predation (both absolutely and relative to its control) under design A when the mimic was common than under design B when the mimic was rare. Judged by its effect on the model, mimicry between the mimic and model is therefore (quasi-)Batesian in nature.

Although the model suffered higher predation when the mimic was common this could be because more attacks ended with taste rejection (i.e. a peck) than with ingestion ('removal'). We calculated the proportion of all model attacks which were pecks. This could be calculated for 71 cases. The general linear model, with design and experiments as factors and day within a block as a covariate, suggested a strong effect of experiment ($F_{1,67} = 1.90$ and $p < 0.001$); thus, the two experiments were analysed separately. In experiment 1, there was an effect of design ($F_{1,30} = 0.44$ and $p < 0.01$) but not day within a block ($F_{1,30} = 0.09$ and $p > 0.2$). Inspection of the data revealed that the effect of design was that the fraction of attacked prey which were pecked but not eaten was higher when the number of mimics was higher (table 1) (mimic common, 79% of attacks are pecks and mimic rare, 66% of attacks are pecks). In experiment 2 there was no effect of design ($F_{1,36} = 1.52$ and $p > 0.2$) or day within a block ($F_{1,36} = 0.24$ and $p > 0.5$) (table 2).

(c) *The effect of mimicry on the mimic*

Under design A, the mimic and model were both equally common at the start of each day. By analysing

Table 1. Percentages of attacks on each prey type which were classified as a ‘peck’: experiment 1

(‘Peck’ approximates to a measure of rejection of a bait before complete ingestion (‘removed’). However, since those which are removed may either have been completely ingested or been partially ingested and taken out of the garden, this distinction is approximate. Condition A, mimic common and mimic control rare and condition B, mimic rare and mimic control common.)

	days 1–10 condition A	days 11–20 condition A	days 21–30 condition B	days 31–40 condition B
model	76.2500	83.6474	65.2778	66.6667
model control	74.6154	80.3571	71.0833	52.7778
mimic	50.7945	55.3429	26.9231	63.3333
mimic control	25.6494	27.1167	48.7473	63.0252
edible control	0.0000	0.0000	0.0000	0.000

how many of each were attacked, we can look for any evidence that the birds were able to discriminate between the two. We calculated the number of mimics attacked on a given day divided by the number of mimics plus models attacked that day. The higher this value is, the more the mimic is attacked relative to models. We fitted a general linear model with experiment (1 or 2) as a fixed factor and day within a block as a covariate. This analysis showed an effect of experiment ($F_{1,37} = 11.6$ and $p < 0.5$), so the analysis was carried out for each experiment separately. For both experiments, there was no effect of day within a block (experiment 1, $F_{1,37} = 1.1$ and $p > 0.25$ and experiment 2, $F_{1,37} = 1.3$ and $p > 0.25$). If the models and mimics were equally likely to be attacked then the mean of our distribution of calculated quotients should be 0.5. This was true for experiment 2 ($t_{19} = 0.10$ and $p > 0.9$), but not for experiment 1 ($t_{19} = 4.75$ and $p < 0.001$) where mimics were attacked more than models. This suggests that the birds were able to differentiate between mimics and models in experiment 1 but not in experiment 2. When they could differentiate, they (unsurprisingly) attacked the less well-defended mimic more than the model.

Although the absolute number of mimics differs between designs A and B, we can standardize for this to some extent by looking at the proportion of the mimic baits available that day which were attacked. We fitted a general linear model with experiment (1 or 2) and design (A or B) as fixed factors and day within a block (days 1–20) as a covariate. There was a significant effect of experiment ($F_{1,75} = 22.3$ and $p < 0.001$); hence, the two experiments were analysed separately. Experiment 1 suggested no effect of design ($F_{1,37} = 1.98$ and $p > 0.1$) but an effect of day ($F_{1,37} = 0.56$ and $p < 0.01$). Visual analysis of the data suggests a gradual reduction in the proportion of mimics attacked during a given design. Design did become significant ($F_{1,29} = 5.9$ and $p < 0.05$) if days 21–30 (when the rare mimic is attacked at unexpectedly high frequency) are deleted. In this case the proportion of mimics attacked was lower when the mimics were rare. This same effect of design was also found in experiment 2 ($F_{1,37} = 84.1$ and $p < 0.001$), as was the trend over time ($F_{1,37} = 6.5$ and $p < 0.05$).

Table 2. Percentages of attacks on each prey type which were classified as a ‘peck’: experiment 2

(‘Peck’ approximates to a measure of rejection of a bait before complete ingestion (‘removed’). However, since those which are removed may either have been completely ingested or been partially ingested and taken out of the garden, this distinction is approximate. Condition A, mimic common and mimic control rare and condition B, mimic rare and mimic control common.)

	days 1–10 condition B	days 11–20 condition B	days 21–30 condition A	days 31–40 condition A
model	37.5000	51.9231	31.6986	27.4554
model control	62.7500	58.8095	36.7706	44.3548
mimic	0.0000	52.3810	30.6633	22.4351
mimic control	45.9064	19.3220	1.0000	1.0000
edible control	39.2518	0.5000	0.0000	0.0000

(d) Attacks on controls

A general linear model with the number of model controls attacked as a dependent variable suggested a strong effect of experiment ($F_{1,75} = 14.5$ and $p < 0.001$); hence, the two experiments were analysed separately. Experiment 1 showed no effect of either design ($F_{1,37} = 0.46$ and $p > 0.5$) or day within a block ($F_{1,37} = 2.6$ and $p > 0.1$): model controls were attacked at a low, constant rate. In contrast, both design ($F_{1,37} = 33.0$ and $p < 0.001$) and day within a block ($F_{1,37} = 5.0$ and $p < 0.05$) were significant in experiment 2. If we assume that learning was still occurring in the first ten days and delete these from the analysis, then design no longer shows a significant effect ($F_{1,27} = 3.3$ and $p > 0.05$) but day within a block is still significant ($F_{1,27} = 22.2$ and $p < 0.001$). However, it is clear that, during days 21–30, there was an unexpectedly high rate of attacks on model controls. During days 21–40 there was an overall reduction in the attack rates on model controls, which suggests learning by naive predators. We therefore suspect that the rise in attacks on model controls seen between designs B and A resulted from an influx of some naive predators.

Although the absolute number of mimic controls differed between designs A and B, we can standardize for this to some extent by looking at the proportion of the mimic control baits available that day which were attacked. We fitted a general linear model with experiment (1 or 2) and design (A or B) as fixed factors and day within a block (days 1–20) as a covariate. There was no significant effect of experiment ($F_{1,75} = 0.01$ and $p > 0.5$) or day within a block ($F_{1,75} = 0.01$ and $p > 0.5$), but there was an effect of design ($F_{1,75} = 5.17$ and $p < 0.001$). Visual inspection of the data showed that the proportion of mimic control baits attacked was much higher when they were rare (and the mimic was common): on average 90% were attacked when five baits were placed out each day and only 40% when 20 were available. Increasing the numbers of mimic controls (from condition B to A) is analogous to adding a moderately defended mimic to an identical but rare mimic. When this happens there is clearly a Müllerian effect in that the survival rate of both co-mimics increases.

One surprising aspect of the results is the fact that, in experiment 1, the common mimic control (days 21–40) was attacked at a generally lower rate than the common mimic (days 1–20). There may have been some change in the population of visiting birds during the length of the experiment which explains this anomaly.

(e) *Attacks on edible controls*

The number of edible controls attacked each day was independent of experiment ($F_{1,75} = 0.91$ and $p > 0.3$) and design ($F_{1,75} = 0.91$ and $p > 0.3$) but was affected by day within a block ($F_{1,75} = 23.6$ and $p < 0.001$). Visual inspection of the data showed that this effect is an increase in the attack rate on these items over time.

4. DISCUSSION

We have presented the first experiments designed to test the key predictions of conventional and unconventional theories of Müllerian mimicry. They show that moderately defended artificial prey can act either as Müllerian mimics or as parasitic quasi-Batesian mimics (diluting the protection of better defended prey). Conventional theories have failed to accommodate these results (Müller 1879; Turner *et al.* 1984; Mallet & Joron 1999); instead, they predict Müllerian mimicry in both cases. In contrast, the results show that predators can behave in ways qualitatively predicted by Pavlovian learning theory (Speed 1993a) and, thus, match the predictions of unconventional, quasi-Batesian theories (reviewed in Speed & Turner 1999b).

However, the moderately defended mimic is categorically not an edible Batesian mimic. As its non-mimetic control shows, its aversive chemicals clearly confer defence: when mimic control was common and ignoring periods of learning, it had considerable protection compared to the edible control prey (31% were attacked during days 31–40 in experiment 1 and 54.5% were attacked during days 11–20 in experiment 2). During this time all edible controls were attacked and eaten. In addition, when a mimic control was common its survival rate was much higher than when it was rare: this is typical of a defended rather than an edible prey (e.g. Turner *et al.* 1984; Speed 1993a). In this situation a mimic control clearly functions as a Müllerian co-mimic to similarly defended prey. It is also easy to predict that a mimic control could function as a Batesian model to an edible bait.

Mallet & Joron (1999) (and Joron & Mallet 1998) argued that quasi-Batesian mimicry must be very rare, if it exists at all, because intermediate levels of defence occur very rarely and most prey will be completely accepted or rejected by predators. Joron & Mallet (1998) cited only one paper as supporting evidence (Turner *et al.* 1984; data originally from Morell & Turner (1970)). However, Morell & Turner (1970) soaked their unpalatable prey in a 75% quinine solution in order to produce 'extremely distasteful' prey. No attempt was made to generate intermediate levels of defence (see Lea & Turner (1971) who did). Indeed, it is quite easy to obtain stable intermediate attack rates as our results show (see also Brower (1958, p. 37), with data pooled across birds E1–4) (Platt *et al.* 1971; Gittleman *et al.* 1980; Grieg-Smith 1987; Speed 1990). In an extensive recent study, Sargent (1995) found a full spectrum of

acceptabilities in lepidoptera, in which only 11% fell into the range of moderate–high unpalatability, with 33% being slightly acceptable–unacceptable. Joron & Mallet's (1998) hypothesis of complete acceptance or complete rejection does not have strong empirical support and, hence, refutations of quasi-Batesian mimicry on these grounds appear to be unfounded.

The results of our experiments question the behavioural assumptions made in recent critiques of unconventional theories of Müllerian mimicry. The traditional Müllerian framework considers defended prey in terms of their toxin content (Turner *et al.* 1984; Mallet & Joron 1999), but not their nutritional components. It assumes that, after learning about any defended prey, avoidance will be complete and that the addition of less well-defended mimics to a well-defended species can only therefore add protection. However, it seems more likely that the birds used knowledge from Pavlovian learning (Speed 1993a; MacDougall & Dawkins 1998) in their foraging decisions, trading off the unpleasantness of quinine and mustard against the nutritional benefits which the baits provided. The moderately defended mimics in this experiment therefore diluted protection of their better defended models because of their greater nutritional use (see also Srygley & Kingsolver 1998).

Experiments with free-living garden birds and pastry prey provide a very useful way of investigating real ecological systems. However, two important limitations should be noted. First, potential variability in the population of garden birds over time may affect the attack rates. It appears to have generated some anomalies in our results (see, for example, the model control data in experiment 2). Second, artificial baits rather than real insects have been used as prey and, hence, the trade-off of nutrition versus aversion experienced by the birds is different to that usually encountered. Data from carefully controlled laboratory studies and from populations of individually identifiable predators are essential for the decisive testing of Müllerian mimicry theory. Despite these reservations, competing theories of Müllerian mimicry have produced opposing predictions which have been relatively easy to test. Our results fail to support conventional theories of mimicry, but can be explained by unconventional theories which assume that predators learn in a Pavlovian fashion. We suggest that these experiments represent the beginning of a resolution to this old controversy.

Valued advice on the manuscript was given by Nicola Marples, Mike Weale and an anonymous referee. M.S. thanks Jim Mallet and John Turner for numerous enjoyable discussions and Paul Slater and Lori Munro for pilot work. M.S. was funded by the Liverpool Hope Research Fund and a Sheppard Worlock research fellowship. Graeme Ruxton acknowledges support from the Natural Environment Research Council and the Nuffield Foundation.

REFERENCES

- Alcock, J. 1970 Punishment levels and the response of black-capped chickadees (*Parus atricapillus*) to three kinds of artificial seeds. *Anim. Behav.* **18**, 592–599.
- Brower, J. V. Z. 1958 Experimental studies of mimicry in some North American butterflies. I. The monarch, *Danaus plexippus* and viceroy, *Limenitis archippus archippus*. *Evolution* **12**, 32–47.

- Brower, L. P., Brower, J. V. Z. & Collins, C. T. 1963 Experimental studies of mimicry. VII. Relative palatability and Müllerian mimicry among neotropical butterflies of the subfamily Heliconiinae. *Zoologica* **48**, 65–84.
- Endler, J. A. 1991 Interactions between predators and prey. In *Behavioural ecology* (ed. J. R. Krebs & N. B. Davies), pp. 169–196. Oxford, UK: Blackwell.
- Gittleman, J. L., Harvey, P. H. & Greenwood, P. J. 1980 The evolution of conspicuous colouration: some experiments in bad taste. *Anim. Behav.* **28**, 879–899.
- Grieg-Smith, P. W. 1987 Aversions of starlings and sparrows to unfamiliar, unexpected or unusual flavours and colours in food. *Ethology* **74**, 155–163.
- Huheey, J. E. 1976 Studies in warning coloration and mimicry. VII. Evolutionary consequences of Batesian–Müllerian spectrum: a model for Müllerian mimicry. *Evolution* **30**, 86–93.
- Huheey, J. E. 1988 Mathematical models of mimicry. *Am. Nat.* **131**, S22–S41.
- Ikin, M. & Turner, J. R. G. 1972 Experiments on mimicry: *gestalt* perception and the evolution of genetic linkage. *Nature* **239**, 525–527.
- Joron, M. & Mallet, J. 1998 Diversity in mimicry: paradox or paradigm. *Trends Ecol. Evol.* **13**, 461–466.
- Lea, G. & Turner, J. R. G. 1971 Experiments on mimicry. II. The effect of a Batesian mimic on its model. *Behaviour* **42**, 131–151.
- MacDougall, A. & Dawkins, M. S. 1998 Predator discrimination error and the benefits of Müllerian mimicry. *Anim. Behav.* **55**, 1281–1288.
- Mallet, J. & Joron, M. 1999 The evolution of diversity in warning colour and mimicry. *A. Rev. Ecol. Syst.* **30**, 201–203.
- Marples, N. M. 1993 Do wild birds distinguish palatable and unpalatable prey types? *Anim. Behav.* **46**, 347–354.
- Marples, N. M., Roper, T. J. & Harper, D. G. C. 1998 Responses of wild birds to novel prey: evidence of dietary conservatism. *Oikos* **83**, 161–165.
- Morrell, G. M. & Turner, J. R. G. 1970 Experiments on mimicry. I. The responses of wild birds to artificial prey. *Behaviour* **36**, 116–130.
- Müller, F. 1879 Ituna and Thyridia; a remarkable case of mimicry in butterflies. *Trans. Melod. R. Entomol. Soc. Lond.* **1879**, xx–xxix.
- Owen, R. E. & Owen, A. R. G. 1984 Mathematical paradigms for mimicry: recurrent sampling. *J. Theor. Biol.* **109**, 217–247.
- Platt, A. P., Coppinger, R. P. & Brower, L. P. 1971 Demonstration of the selective advantage of mimetic *Limenitis* butterflies presented to caged avian predators. *Evolution* **25**, 692–701.
- Ruxton, G. D. 1998 Wolves in sheep's clothing. *Nature* **394**, 833–834.
- Sargent, T. D. 1995 On the relative acceptabilities of local butterflies and moths to local birds. *J. Lepidopt. Soc.* **49**, 148–162.
- Speed, M. P. 1990 Mimicry and the psychology of predation. PhD thesis, University of Leeds, UK.
- Speed, M. P. 1993a Müllerian mimicry and the psychology of predation. *Anim. Behav.* **45**, 571–580.
- Speed, M. P. 1993b When is mimicry good for predators? *Anim. Behav.* **46**, 1246–1247.
- Speed, M. P. 1998 Mistakes not necessary for Müllerian mimicry. *Nature* **396**, 323.
- Speed, M. P. 1999 Robot predators in virtual ecologies: the importance of memory in mimicry studies. *Anim. Behav.* **57**, 203–213.
- Speed, M. P. & Turner, J. R. G. 1999a Virtual predators, receiver psychology, and doubts about Müllerian mimicry: comments on MacDougall and Dawkins. *Anim. Behav.* **58**, F10–F13.
- Speed, M. P. & Turner, J. R. G. 1999b Learning and memory in mimicry. II. Do we understand the mimicry spectrum? *Biol. J. Linn. Soc.* **67**, 281–312.
- Srygley, R. B. & Kingsolver, J. G. 1998 Red-wing blackbird reproductive behaviour and the palatability, flight performance and morphology of temperate pierid butterflies (*Colias*, *Pieris* and *Pontia*). *Biol. J. Linn. Soc.* **64**, 41–55.
- Turner, J. R. G. 1971 Studies of Müllerian mimicry and its evolution in burnet moths and heliconid butterflies. In *Ecological genetics and evolution* (ed. E. R. Creed), pp. 224–260. Oxford, UK: Blackwell.
- Turner, J. R. G. 1987 The evolutionary dynamics of Batesian and Müllerian mimicry: similarities and differences. *Ecol. Entomol.* **12**, 81–95.
- Turner, J. R. G. & Speed, M. P. 1996 Learning and memory in mimicry. I. Simulations of laboratory experiments. *Phil. Trans. R. Soc. Lond.* **B351**, 1157–1170.
- Turner, J. R. G., Kearney, E. P. & Exton, L. S. 1984 Mimicry and the Monte Carlo predator: the palatability spectrum and the origins of mimicry. *Biol. J. Linn. Soc.* **23**, 247–268.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.